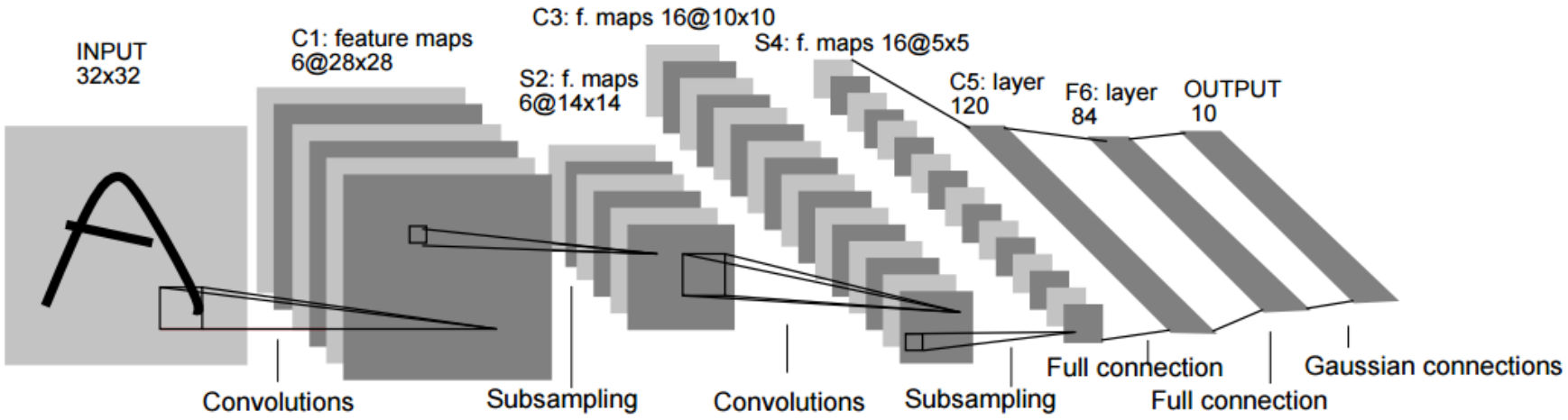
Q1-1: Select the correct option about LeNet-5.

A. *LeNet-5 architecture has subsampling layers which essentially does pooling operation.*

B. *Fully Connected Network is used in the end to obtain softmax scores.*

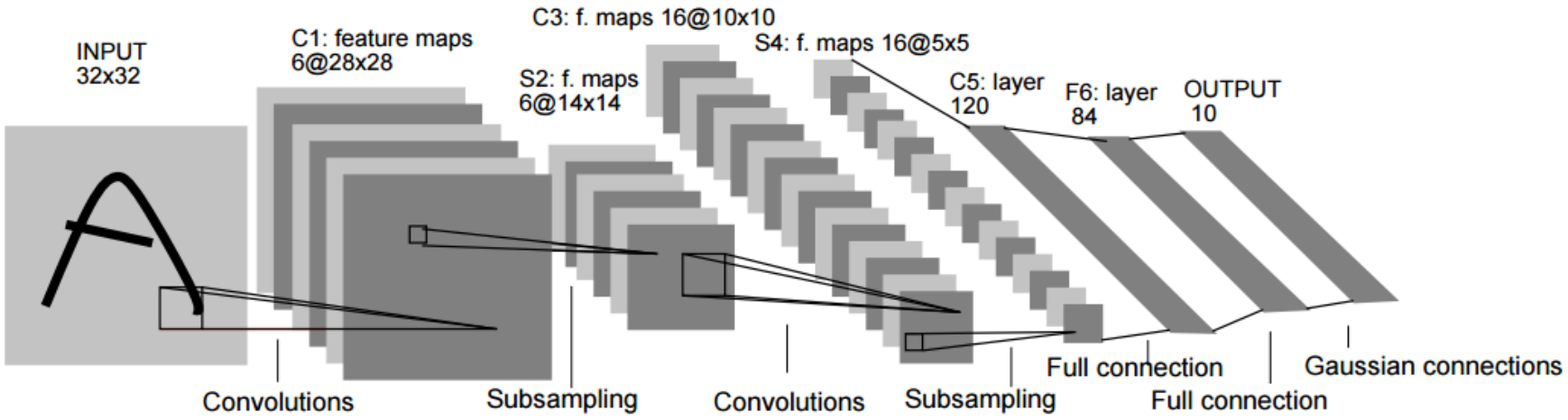


1. Both statements are true.
2. Both statements are false.
3. Statement A is true, Statement B is false.
4. Statement B is true, Statement A is false.

Q1-1: Select the correct option about LeNet-5.

A. *LeNet-5 architecture has subsampling layers which essentially does pooling operation.*

B. *Fully Connected Network is used in the end to obtain softmax scores.*



1. **Both statements are true.** 
2. Both statements are false.
3. Statement A is true, Statement B is false.
4. Statement B is true, Statement A is false.

Q2-1: Are these statements true or false?
(A) Order matters in sequential data.
(B) A batch of sequential data always contains sequences of a same length.

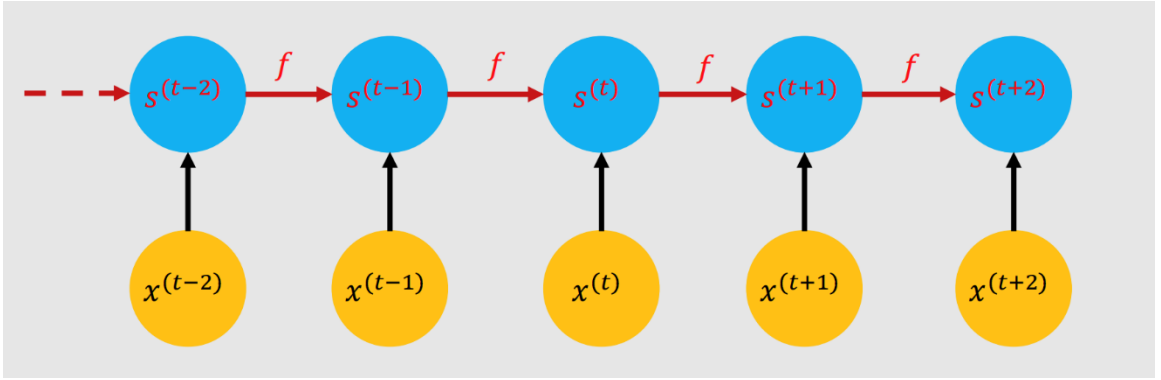
1. True, True
2. True, False
3. False, True
4. False, False

Q2-1: Are these statements true or false?
(A) Order matters in sequential data.
(B) A batch of sequential data always contains sequences of a same length.

1. True, True
2. **True, False** 
3. False, True
4. False, False

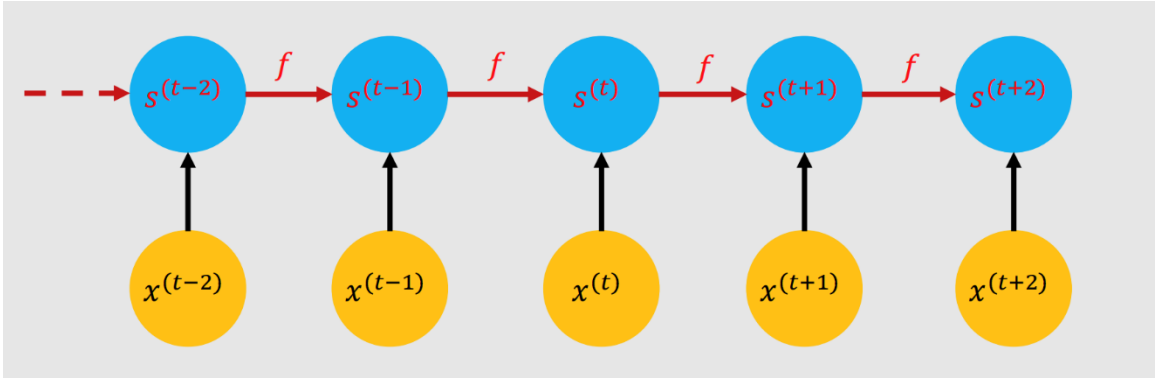
(A) As is shown by its name “sequential”, order matters in sequential data.
(B) A batch of sequential data can have different length, such as different sentences.

Q2-2: Please choose the representation of $s^{(t+2)}$ in terms of $s^{(t)}, x^{(t)}, x^{(t+1)}, x^{(t+2)}$ in the following dynamic system $s^{(t+1)} = f_\theta(s^{(t)}, x^{(t+1)})$.



1. $f_\theta(s^{(t)}, x^{(t+1)})$
2. $f_\theta(s^{(t)}, x^{(t+2)})$
3. $f_\theta(f_\theta(s^{(t)}, x^{(t)}), x^{(t+1)})$
4. $f_\theta(f_\theta(s^{(t)}, x^{(t+1)}), x^{(t+2)})$

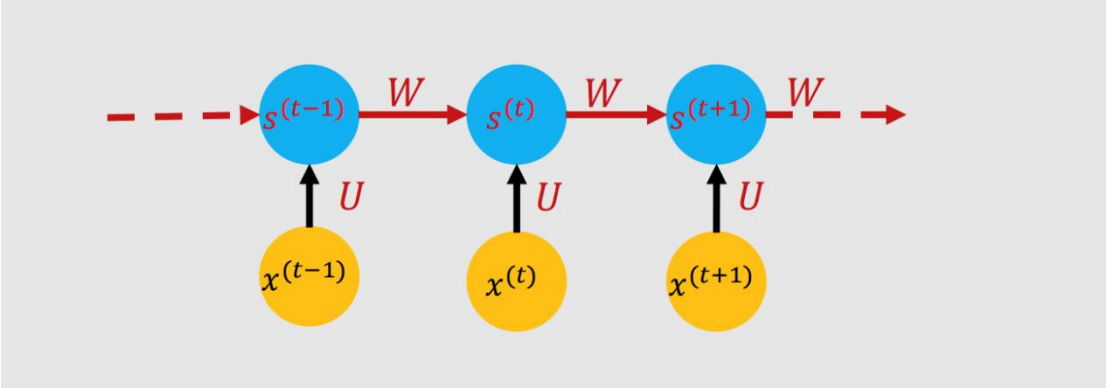
Q2-2: Please choose the representation of $s^{(t+2)}$ in terms of $s^{(t)}, x^{(t)}, x^{(t+1)}, x^{(t+2)}$ in the following dynamic system $s^{(t+1)} = f_\theta(s^{(t)}, x^{(t+1)})$.



1. $f_\theta(s^{(t)}, x^{(t+1)})$
2. $f_\theta(s^{(t)}, x^{(t+2)})$
3. $f_\theta(f_\theta(s^{(t)}, x^{(t)}), x^{(t+1)})$
4. $\boldsymbol{f_\theta(f_\theta(s^{(t)}, x^{(t+1)}), x^{(t+2)})}$ ⬅

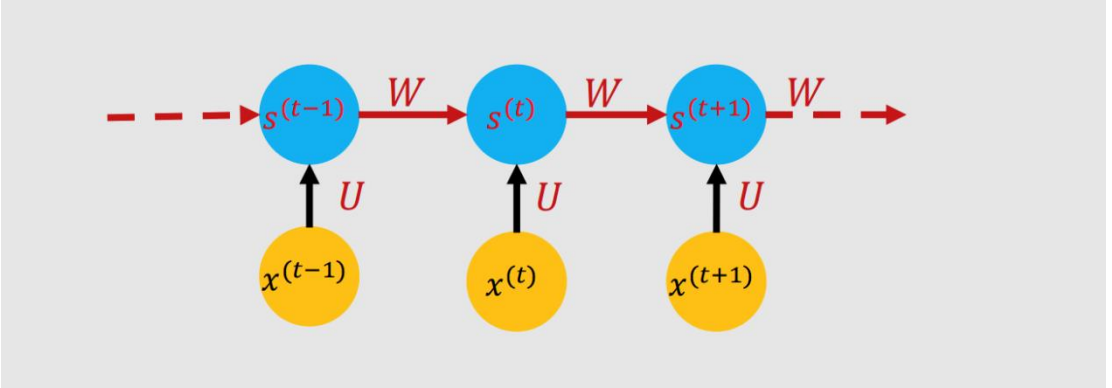
As is shown in this dynamic system, we have $s^{(t+2)} = f_\theta\left(s^{(t+1)}, x^{(t+2)}\right) = f_\theta\left(f_\theta\left(s^{(t)}, x^{(t+1)}\right), x^{(t+2)}\right)$, as $s^{(t+1)} = f_\theta\left(s^{(t)}, x^{(t+1)}\right)$.

Q2-3: Are these statements true or false?
(A) The hidden state $s^{(t)}$ is the linear combination of the previous hidden state $s^{(t-1)}$ and the external data $x^{(t)}$.
(B) Sharing functions and parameters in RNN leads to inherent limitation on the learning ability of the model.



1. True, True
2. True, False
3. False, True
4. False, False

Q2-3: Are these statements true or false?
(A) The hidden state $s^{(t)}$ is the linear combination of the previous hidden state $s^{(t-1)}$ and the external data $x^{(t)}$.
(B) Sharing functions and parameters in RNN leads to inherent limitation on the learning ability of the model.



1. True, True
2. True, False
3. False, True
4. **False, False** ⟵

(A) We need to use an activation function to compute the hidden states, so it's not linear.
(B) As is shown in the lecture, such RNN of a finite size can be universal.